Independent Research Project

Prediction of wildfire duration and final burned area with image-based Machine learning

Author name: Hansong Xiao

Supervisor: Dr. Sibo Cheng, Dr. Rossella Arcucci

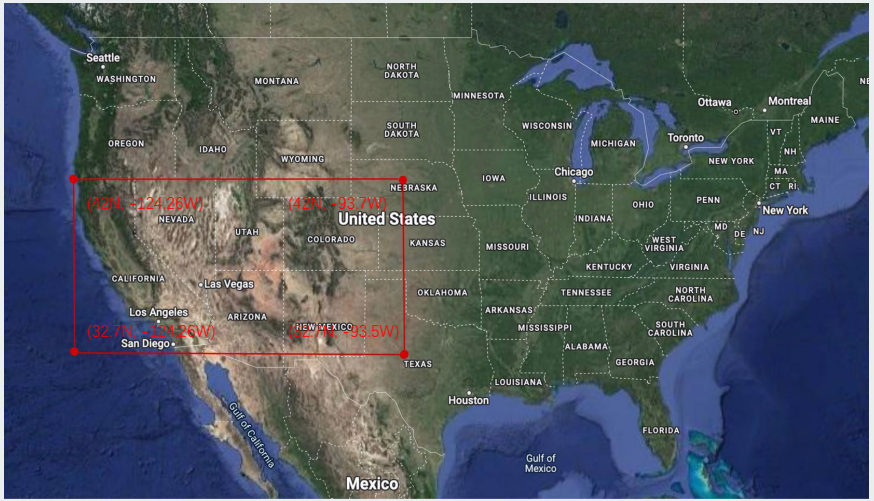
Introduction

- Detriment of Wildfire
  - Economy Loss
    - House Price Depreciate
  - Ecosystem
    - Burning Grouped Animal
    - More than 4 million km^2 burned area

Introduction

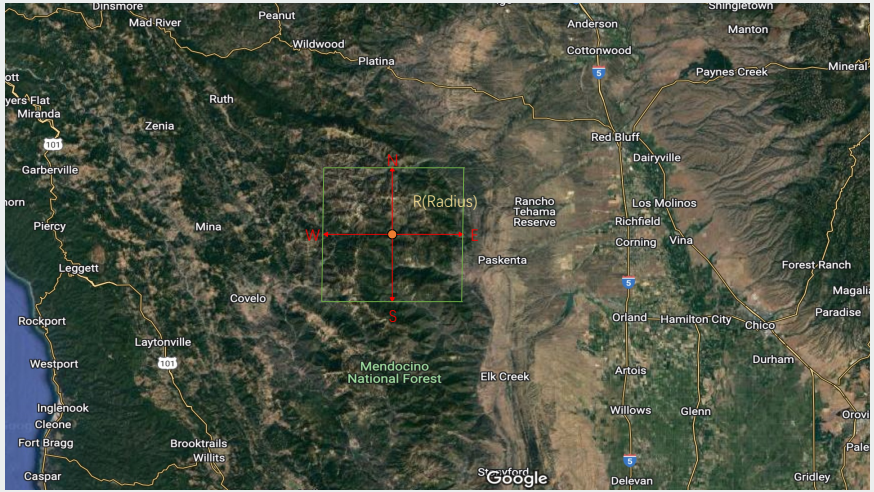
- Traditional Method (Grid-Based Model)
  - Method Name
    - Computational Fluid Dynamics (CFD)
    - Celluar Automata
  - Disadvantage of Traditional Way
    - Time-Consuming
    - Difficult to predict ahead of a fire
- Fast-Decision Method
  - Goals
    - Making prediction after fire
    - Focusing on the duration of fire
  - Method Type
    - Regression ML Model
    - Image-Based ML Model

Data Gathering



- Reason for target area
  - Similarity in landscape
  - Number of wildfire
  - Representative of economic loss

Data Gathering



- Expand every fire with a Radius
- Extract different features from different database

Dataset	Extracted Feature
Global Fire Atlas with Characteristics of Individual Fires	fires location, duration
Global ALOS mTPI	Landform Slope
Copernicus Global Land Cover Layers	Landform cover map
ERA5 Monthly Aggregates - Latest Climate Reanalysis	Wind data

Methodology (Data Preprocessing and Standardization)

- Regression ML Model
  - Average value
  - Min-Max Normalization

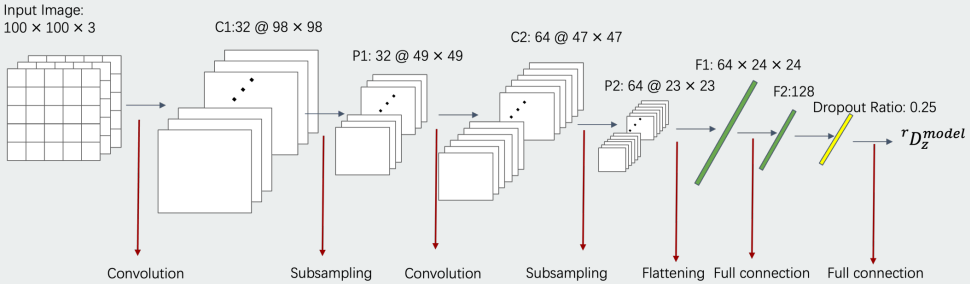
$$X = \frac{X - \min(X)}{\max(X) - \min(X)}$$

- Image-Based Model
  - Cropped or Resized
  - Min-Max Normalization

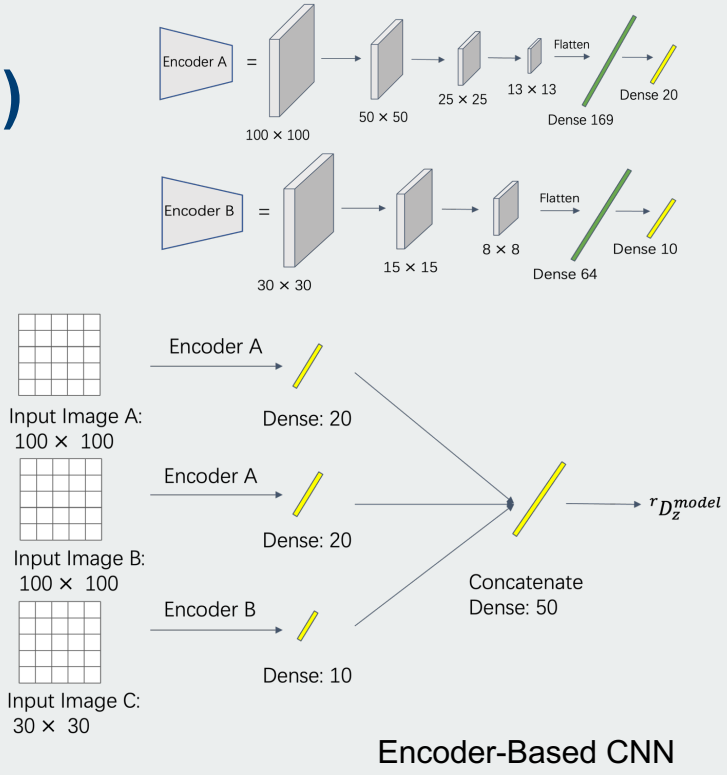
Methodology

- Regression ML Model
  - Random Forest (RF)
    - Extension of bagging algorithm
    - A series of decision trees
  - K-nearest neighbors (KNN)
    - Supervised ML algorithm
    - Averaging nearest point
  - Extrame Gradient Boosting (XGBoost)
    - Preventing Over-fitting
- Image-Based ML Model
  - Multi-layer CNN
  - Encoder-Based CNN

Methodology (Model Structure)



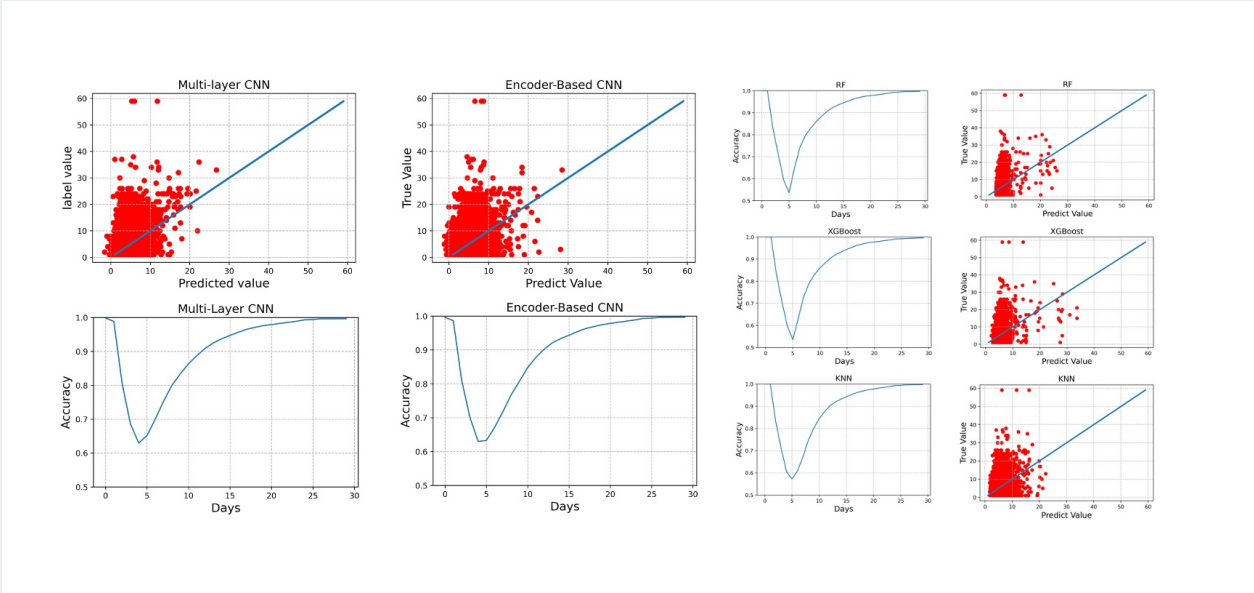
Multi-layer CNN Structure



Encoder-Based CNN

Results

- Prediction Plot
- Accuracy Plot



$$\text{Accuracy} = \frac{TP}{TP+FP}$$

Results & Analysis

Model	RMSE	R2 Score	Accuracy
RF	4.99	0.073	0.536
KNN	5.26	-0.031	0.572
XGBoost	5.05	0.049	0.535
Multi-layer CNN	5.05	0.051	0.629
Encoder Based CNN	5.16	0.005	0.631

- Similar RMSE
- Low R2 Score value
- Better Accuracy on Image-Based model

Future Improvement

- More landform feature
- Find a better statistical coefficient
- Find better model structure to improve accuracy

THANKS

Q & A